

DATA SEARCH SYSTEM AND METHOD FOR DISPLAYING DATA SEARCH RESULTS

5 FIELD OF THE INVENTION

The present invention relates, in general, to a data search system for managing, as sets, data search results based on search criteria, and, more specifically, to a user interface scheme for visually representing data search results.

10

BACKGROUND OF THE INVENTION

Results of data search through a database may be represented as a set according to each search criterion, and the size or magnitude of the set is often represented as a numerical value. The user may use this numerical
15 value to intuitively understand the size of the set. However, as there is more than one search criterion and the resulting sets become complicated accordingly, it could become difficult to recognize the relationship in size or magnitude among those sets, relevancy of the search criteria, and so on, merely by way of numerical values.

20 In order to provide improvement on this aspect, data management systems are traditionally available such that the size of the set is visually represented. They include, for example, a data management system that employs a technique by which a set is represented by a rectangular or circular Venn diagram determined by calculation of a geometric formula.

25 However, even with such a data management system, if multiple Venn diagrams exist, it is difficult to display them within the same display region, while maintaining a relative relationship of each Venn diagram with one another, and it may become impossible to do so depending on the search criteria or the size of the set or product set. For example, it is impossible to
30 display three sets as shown in FIG. 8 (sets A, B, and C) within the same display region as rectangular or circular Venn diagrams, while retaining a relative relationship of each set.

Additionally, there may be cases where in addition to Venn diagrams created under certain search criteria, one may want to create new Venn diagrams with other criteria, such as time and territory, partially added and compare them with one another, but such processing could not be achieved with conventional data management systems.

Thus, it is an objective of the present invention to provide a data search system and a method for displaying data search results that can overcome such shortcomings.

It is another objective of the present invention to provide a computer program and storage medium that may allow a computer to function as such a data search system.

SUMMARY OF THE INVENTION

A data search system according to the present invention comprises a data acquisition means for acquiring results of a data search based on certain search criteria; an image processing means for representing, as a set for each of the search criteria, the results of the data search acquired by the data acquisition means and converting each set to an n-polygonal image having an area corresponding to the magnitude of element count of each set; and a display control means for causing each n-polygonal image converted by the image processing means to be displayed on a predetermined display device, wherein each of the multiple n-polygonal images is displayed within the same display region while maintaining a relative relationship with other n-polygonal images based on the search criteria.

An element of a set is an element for quantitatively representing the magnitude of the set, such as, for example, a numerical value "n" is a natural number, which is usually 5 or greater, and, more preferably, about 12 to 24. This value is a numerical value that may be changed a posteriori, depending on the magnitude of the set, the number of other n-polygonal images, or their relative relationships.

To enable a relationship among search criteria to be visually recognized in accordance with the display positions of the n-polygonal images, each of the sets is converted to an n-polygonal image having an area

corresponding to the magnitude of element count of each set and a relative relationship in position based on the search criteria. A result of a data search obtained by a Boolean AND operation of search criteria is represented by an overlap of the n-polygonal images converted from the respective sets for the search criteria.

The shape and area of an n-polygon that is predicated on the n-polygonal image may be identified by arithmetic operation. For example, it is derived by repeating a convergent calculation for drawing under certain restricting parameters. In that case, the convergent calculation is initiated under restricting parameters where a convex polygon appears, and the restricting parameters are switched to such parameters that a concave polygon appears if it does not converge as a convex polygon even after it is repeated a predetermined number of times.

The image processing means identifies the shape of the n-polygonal image (or position of each vertex) by use of a radius and argument of polar coordinates, and calculates the area, S_i , of the n-polygonal image by use of an operational expression having the radius, r_i , and n , as its variables. More specifically, assuming that the radius of the polar coordinates is r_i where i is an integer greater than zero and not greater than n), then the argument, θ_i , of the polar coordinates is expressed as $(2\pi/n) \cdot i$, as is well known, so that the area, S_i , of the n-polygonal image may be determined by the following equation:

$$S_i = \frac{1}{2} \sin \frac{2\pi}{n} \sum_{i=0}^{n-1} r_i r_{i+1} \dots (1)$$

If a plurality of sets contain any product set determined by a Boolean AND operation, the area S_i of that product set is derived as a result of convergent calculation for increase or decrease in the radius r_i .

For applications where changes in time frame for certain search criteria are to be ascertained, the data search system further comprises an image storage means for storing n-polygonal images converted as the data search

results where only a portion of the search criteria differs, according to the search criteria that are commonly shared by the rest. In that case, the image processing means causes multiple n-polygonal images that vary with changes in a portion of the search criteria to be displayed within the same display region.

As described above, if the shape of the n-polygonal image is identified by the radius γ_i and argument θ_i of the polar coordinates and the area S_i of that n-polygonal image is calculated by the operational expression having the radius γ_i and n as its variables, the image processing means places in the same position the poles of said polar coordinates and their initial lines for a plurality of n-polygonal images displayed within the same display region.

The method of present invention includes a method for displaying a data search result is such that a result of data search based on certain search criteria is represented as a set according to each of said search criteria, and each set is converted to an n-polygonal image having an area corresponding to the magnitude of element count of each set and a relative relationship in position among the search criteria, so that when each n-polygonal image converted is displayed on a predetermined display device, each of the multiple n-polygonal images is displayed within the same display region, while maintaining a relative relationship with other n-polygonal images based on the search criteria.

The present invention also provides a computer program in a computer that conducts a data search based on certain search criteria. The said computer comprises a data acquisition means for acquiring results of data search based on the search criteria; an image processing means for representing the search results acquired by the data acquisition means as sets according to each of the search criteria, and converting each set to an n-polygonal image having an area corresponding to the magnitude of element count of each set and a relative relationship in position among the search criteria; and a display control means for causing a predetermined display device to display each of the n-polygonal images converted by the image processing means, wherein the computer program causes each of the n-polygonal images to be displayed within the same display region while

5

10

15

20

25

30

5

DETAILED DESCRIPTION OF A PREFERRED EMBODIMENT

An embodiment of the present invention is described below where the mechanism of the present invention is applied to a network-based database management system.

5 With reference to Fig. 1, a database management system comprises a management server 20 that forms an example of the data search system, and a search server 30, both connected to a computer network L. In this embodiment, it is contemplated that the computer network L is the Internet, to which clients 10 manipulated by users and various information servers 40 are
10 connected.

The client 10 is a general-purpose computer terminal having communication capability for conducting bi-directional communication with the Internet L and display capability for presenting contents on a display, and also incorporates a browser. The client 10 makes a request for data search to the
15 management server 20 via the Internet L and presents results of the data search on the browser. To the request for data search are attached search criteria.

The management server 20 and search server 30 are computers having Web server capabilities that are accessible to the Internet L, and
20 implement required functions related to data search and display of search results, in addition to the Web server capabilities, in cooperation with hardware resources and computer programs. More specifically, the search server 30 implements the capability of conducting a data search to the information server 40, and the management server 20 implements functions
25 of a data manager unit 21, an image processing unit 22, an image storage unit 23, and a display control unit 24.

These functions are each realized by cooperation between the computer programs stored in a computer-readable storage medium and the hardware resources of the server that reads them.

30 By way of example, the management server 20 and search server 30 are configured as separate servers, although it will be appreciated that they may be implemented as a single server system.

The data manager unit 21 of the management server 20 stores a HTML page for presentation on the browser of the client 10 and a CGI script for activating the search server. It invokes the CGI script by acquiring a data search request and search criteria from the client 10 via the HTML page, and
5 this CGI script is used to activate the search server. The CGI script obtains from the search server 30 a HTML image that represents the data search result, and temporarily stores it in a buffer area within the management server 20.

The image processing unit 22, which is intended to perform the
10 process of displaying a characteristic data search result, displays, as a set according to each search criterion, a data search result obtained corresponding to a certain search criterion, and converts each set to a Venn diagram that is an n-polygonal image having an area corresponding to the magnitude of element count of each set. Elements of a set may include ones
15 that can represent a relative relationship in magnitude, i.e., a numerical value, such as data count, for example.

For data search results based on Boolean AND operation of search criteria or other logic conditions, they are represented by overlaps of the Venn diagrams converted by the respective sets of the search criteria.

20 Greater detail of the above processing is described herein below.

The image storage unit 23 stores n-polygonal Venn diagrams for the respective search criteria, so that they may be read out using any of the search criteria as a key. When the search time frame, such as the search date, is contained as one of the search criteria, Venn diagrams based on all
25 search criteria by search date are stored. In that case, using the search date as a key, all the relevant Venn diagrams can be read out. Additionally, Venn diagrams where only a portion of the search criteria is different may be stored according to each of the remaining search criteria, so that they may be read out using the remaining search criteria as keys.

30 The display control unit 24 performs control to allow the browser of the client 10 to display each of the Venn diagrams drawn by the image processing unit 23, together with items and so forth for displaying these Venn diagrams.

The procedures executed by the image processing unit 22 are described in detail below.

It is presumed that the magnitude of a set that is a data search result is all represented by a numerical value in the form of hit count. If there are multiple search criteria, with some relative relationship among these search criteria, the resulting data search results may sometimes maintain the above relative relationship with other data search results. In such cases, each of the multiple Venn diagrams is displayed within the same region of the display, while maintaining the relative relationship with other Venn diagrams based on the search results.

This process is specifically performed in accordance with the procedures illustrated in FIG. 2. For the sake of simplicity of explanation, it is assumed that three sets, "Tokyo Residents" (total count of 300 hits), "Male" (total count of 1000 hits), and "Contract Years of Three or Longer" (total count of 800 hits) corresponding to the results of respective data searches based on three search criteria, "Address = 'Tokyo'", "Sex = 'Male'", and "Contract Years >= Three Years" for the customer data on a certain date are obtained as shown in FIG. 3 (a).

In the example of FIG. 3 (a), the left-hand side member is search criteria, while the right-hand side member is a data search result. For example, the top of the list indicates that for customers who are residents in Tokyo, not male, and have contract years of less than three years, there are 300 hits of data. Meanwhile, the bottom of the list indicates that for customers who are residents in Tokyo, male, and have contract years of three years or longer, there are 100 hits of data. The rest of the search criteria and data search results are also seen in a similar manner.

In the present embodiment, with attention paid to the product set that satisfies all the search criteria (in the example of FIG. 3 (a), the bottom of the list, i.e., 100 hits), a polar coordinate system with its pole located in or near the center of the product set is defined (Step S101: Yes, S102). If there is no product set, that is, there is only a single set, a polar coordinate system with its pole located in the center of the set is defined (Step S101: No, S103).

Thus, all the search criteria may be represented by polar coordinates having the pole of the polar coordinate system as their respective pole.

Once the pole of the polar coordinate system is defined, a plurality of coordinate axes passing across the pole are set (Step S104). The coordinate axes are set by segmenting 360 degrees by equal intervals of n . Then, an n -polygonal Venn diagram is represented by a combination of vertexes identified by the radius γ_i and argument θ_i on each coordinate axis (Step S105). If there are multiple Venn diagrams, the relative relationship in position among the Venn diagrams is adjusted.

The greater the value of n , the greater the flexibility of the shape of an individual Venn diagram, so that the relative relationship with other Venn diagrams can be represented more accurately; however, the computational burden is also increased accordingly; thus, an appropriate value should be set in consideration of the number of sets and the magnitude of an individual set. In the present embodiment, "12" is set as the value of n . In this case, the value of i for the radius γ_i and argument θ_i of the polar coordinates ranges from 0 to 12. Note, however, that the initial line (γ_0, θ_0) and end line (γ_n, θ_n) assume the same value.

It will be appreciated that the value of n may be changed a posteriori. That is, if some of the sets have magnitudes that differ extremely from those of the rest, the relative relationship among the Venn diagrams may not sometimes be displayed properly at the same time, in which case by changing the value of n , adjustments can be made a posteriori so that they are easy to view.

The shape of the Venn diagram can be identified by repeating the convergent calculation for the radius γ_i that selectively uses any of the restricting parameters stored in the restrictions storage region (not shown) in the management server 20. This convergent calculation is an arithmetic operation such that the Venn diagram that represents the magnitude (numerical value) of an individual set is converged as a convex polygon; the process, which is preset as a routine, is activated upon entry of a restricting parameter, and automatically executed.

As restricting parameters, various types may be utilized. For example, assuming that a Venn diagram is displayed by a group of pixels on the display of the client 10 and in consideration of a rule of thumb that human visual perception recognizes more than a certain number of pixels, a condition where "a convergence point is such that the difference between the magnitude (actual value) of the set and the area (value obtained by drawing) is 10 pixels or less" may be used as one of the restricting parameters. A parameter for identifying the shape when drawn may also be used as one of the restrictions.

The convergent calculation is initiated under restricting parameters where the Venn diagram is converged to a convex polygon in a certain position, that is because a convex polygon makes it easier to understand the magnitude of a set due to human visual perception. However, although it is not always necessary to do so, a Venn diagram of concave polygonal shape may also be used. Typically, the convergent calculation is initiated under restricting parameters where it is converged to a convex polygon, and if it is not converged as a convex polygon after the convergent calculation is performed a predetermined number of times, the restricting parameters are switched by removing the restrictions one by one, so that a Venn diagram of concave polygonal shape appears.

Once the shape and position of the Venn diagram is determined by the convergent calculation, the area S_i of each Venn diagram is calculated (Step S106). When the set is represented by a dodecagonal Venn diagram by the afore-mentioned polar coordinate system, an individual Venn diagram itself is represented by a combination of twelve triangles. An overlap portion of product sets, if any, determined by a Boolean AND operation is also a dodecagon (a combination of triangles). This is self-evident in terms of algebraic geometry. Thus, the area calculation for Venn diagrams that are data search results can be simplified (sum of the areas of the triangles), and is therefore very convenient.

The area S_i may be specifically derived by performing arithmetic operation on the following equation. This equation suggests that the area of

a dodecagon can be derived by summing the areas of the triangles determined by simple geometric calculation.

$$\begin{aligned} S &= \frac{1}{2} \sin \frac{\pi}{6} \sum_{i=0}^{11} r_i r_{i+1} \\ &= \frac{1}{2} \cdot \frac{1}{2} \sum_{i=0}^{11} r_i r_{i+1} \\ &= \frac{1}{4} \sum_{i=0}^{11} r_i r_{i+1} \end{aligned} \quad \dots (2)$$

5

The areas of product sets and remaining sets may also be derived as results of the convergent calculation for an increase or decrease in the radius γ_i .

When identifying the shape of the Venn diagram, the image processing
10 section 23 confirms the search criteria corresponding to the respective Venn diagrams, and checks if there is any matched search criterion. If there is any matched search criterion, it adds to the Venn diagram the position information such that the poles of the Venn diagrams and their initial lines of the radius coincide with each other (Step S107: Yes, S108). The display control unit 24
15 performs control so that these Venn diagrams are presented on the display of the client 10 (Step S109). If there is no matched search criteria, it performs control so that the Venn diagrams are displayed with their position left unchanged (Step S107: No, S109).

Subsequently, each of the Venn diagrams is stored in the image
20 storage unit 23 according to each of the search criteria, so that they will be read at any time in response to a display request from the client 10 (Step S110).

In this way, on the display of the client 10, three Venn diagrams are displayed within the same display region, while maintaining their relative
25 relationship. That is, while maintaining the condition where the relative hit counts for male customers of 1000 hits, Tokyo resident customers of 700 hits, and over-three-contract-year customers of 800 hits are directly reflected, the

relationship for the respective product sets (male customers who are residents in Tokyo: 100 hits; Tokyo resident customers who have contract years of 3 or longer: 200 hits; male customers who have contract years of 3 or longer: 300 hits; and male customers who are residents in Tokyo and have contract years of 3 or longer: 100 hits) is simultaneously presented on the display.

On the right-hand side in FIG. 3 (b), there are shown date select buttons as an example of time-frame element. This indicates that for the same search criteria, data search results where only a temporal (time frame) element is different are stored in the image storage unit 23. When one of the buttons is selected from the client 10, the customers status on that date is presented under condition where the relative relationship in position of the Venn diagrams is maintained, that is, in the same coordinate system.

FIG. 4 depicts the three Venn diagrams shown in FIG. 3 (b), along with the respective hit data count and the afore-described coordinate axes. It can be seen from this figure that each of the Venn diagrams, including portions overlapped by Boolean AND operation, is also represented by a combination of triangles. By allowing the data count to be displayed, the area of the Venn diagram permits you to intuitively understand the relative relationship with other Venn diagrams, while confirming the correct hit count.

FIG. 5 (a) depicts the past data search results obtained by the same search criteria as in FIG. 3 (a), and FIG. 5 (b) depicts three Venn diagrams created and stored in a similar manner as in FIG. 3 (b), with different time frames selected by the date select buttons located on the right-hand side of the figure. FIG. 6 depicts the Venn diagrams of FIG. 5 (b), along with the count of hit data and the afore-described coordinate axes, where the status of customers on the date of October 10, 1996 is presented.

It will be appreciated that Venn diagrams on all dates may not only be selectively presented in the same coordinate system as shown in FIG. 3 (b) and FIG. 5 (b), but Venn diagrams on different dates may also be presented side by side simultaneously in the same coordinate system. In the latter case, the Venn diagrams on individual dates stored in the image storage unit 23

may be presented in a regularly staggered manner, or simultaneously presented in reduced view.

In this way, with the database management system according to the present embodiment, the image processing unit 22 of the management server
5 20 represents the data search results as sets by search criteria, and converts each set to a dodecagonal Venn diagram having an area corresponding to the magnitude of element count of each set; and the display control unit 24 causes each of the Venn diagrams to be displayed within the same display region, while maintaining the relative relationship with other Venn diagrams
10 based on the search criteria. As such, the human operator who manipulates the client 10 can intuitively understand the data search results.

Furthermore, because the image processing unit 22 permits the number of vertexes (angles) of the Venn diagram to be appropriately changed a posteriori, various display formats for data search results can be achieved,
15 which could not otherwise be realized. For example, FIG. 7 shows an example where the search criteria and data search results shown in FIG. 8 are presented on the display of the client 10, while maintaining the relationship in magnitude of the product sets. Such Venn diagrams could not be expressed with prior art methods that represented Venn diagrams by use
20 of rectangles or circles.

In the present embodiment, a description has been given using a network-based database management system by way of example, although the present invention is not limited to such network types, but implemented as a stand-alone system. In that case, the display means, which corresponds to
25 the display of the client 10, is a display device that comes standard with the stand-alone system.

Networks, if used, are not limited to the Internet, but a local network, such as the intranet, may also be applicable.

In the present embodiment, a description has been given using a
30 dodecagonal Venn diagram as an n-polygonal image by way of example, although the present invention may be implemented with any shape of image that visually represents data search results for multiple search criteria and

permits relative comparison among the respective results, so it is not necessarily restricted to Venn diagrams as in the present embodiment.

As is clear from the above description, according to the present invention, a set that is a result of data search is represented as an n-
5 polygonal image having an area corresponding to its magnitude, so that an easy-to-understand user interface for human operators can be realized.

It should be understood that the invention is not limited to the exact embodiment or construction that has been illustrated and described but that various changes may be made without departing from the spirit and the scope
10 of the invention.